Smartphone-Based Indoor Localization With Integrated Fingerprint Signal

PEIHAO LI[1], XU YANG[1], YUQING YIN[1], SHOUWAN GAO[1,2], AND QIANG NIU[1,2]

[1]School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China
[2]Mine Digitization Engineering Research Center of Minstry of Education of the People's Republic of China, China University of Mining and Technology, Xuzhou 221116, China

Corresponding author: Qiang Niu (niuq@cumt.edu.cn)

ABSTRACT Indoor localization of smartphones has received much attention recently and the smartphone localization is essential to a wide range of applications in office buildings, nursing homes, parking lots, and other public places. Existing solutions relying on inertial sensors or received signal strength suffer from large location errors and poor stability. We observe an opportunity in the recent trend of increasing numbers of wireless transmitters installed in indoor spaces to design a precise and robust indoor localization solution. We can extract fine-grained channel state information from wireless transmitters for indoor fingerprint localization. However, the accuracy of localization relying on a single physical quantity is limited and difficult to self-correct. This study proposes an integrated channel state information (CSI) and magnetic field strength (MFS) localization method (CSMS) that achieves sub-meter accuracy for smartphones. CSMS constructs an integrated fingerprint map of CSI and MFS and proposes the Local Dynamic Time Warping algorithm for geomagnetic tracking and the Multi-Module Data k-Nearest Neighbor algorithm for fusion fingerprint dynamic weighted comparison. By doing so, CSMS outputs enhanced accuracy with low cost, while overcoming the respective drawbacks of each individual sub-system. We conduct extensive experiments in two scenarios to validate the performance of CSMS. The results of experimental show that the mean distance error in both scenarios is less than 0.5m which is significantly superior to existing smartphone-based indoor positioning methods.

INDEX TERMS Indoor localization, smartphone, magnetic fields, channel state information.

I. INTRODUCTION

Accurate indoor localization is a key enabler for many applications on the horizon, such as positioning and navigation in parking lots, real-time monitoring of the elderly in nursing homes, and hazard warnings in construction sites. Due to the ubiquity of indoor localization requirements, the high cost of deployment may lead to it difficult to popularize while the diversity of indoor scene requires the indoor localization to be reliable and easy to expand.

Existing methods to achieve sub-meter localization, such as ZigBee [1] and Ultra-wide Band (UWB) [2], require additional relatively expensive hardware facilities. RFID-based localization [3] has a small coverage and poor stability. Ultrasonic localization [4] attenuates obviously and easy to occur scattering in indoor scenes. None of these methods is

The associate editor coordinating the review of this manuscript and approving it for publication was Di Zhang.

available on smartphones. Localization based on MFS [5] and CSI [6], [7] has better universality in practical application, but the localization accuracy of these methods is not high. As we all know, the magnetic field is susceptible to external interference, and pure magnetic field localization has the possibility of failure. As a fine-grained value of the physical layer, CSI [8] can better reflect the multipath effect in the environment. However, with the increase of the localization range, the localization accuracy of CSI significantly declines. Besides, currently, computers are widely utilized to collect CSI data with poor mobility.

Nowadays, wireless transmitters are ubiquitous in all areas of the society, such as office buildings, museums and shopping malls. In most cases, we can receive multiple Wi-Fi signals [9], [10] in the same location, and the location of signal source usually does not change. At the same time, due to the interference of the reinforced concrete structure of the building to the geomagnetic field in the indoor scenes,

the magnetic field in local area presents a unique distribution, and very stable [5]. We believe that CSI extracted from multiple Wi-Fi signals can be utilized together with magnetic field strength as fingerprint for localization to improve accuracy.

However, translating this initial idea into a specific localization system presents a number of challenges. First, the platform of data collection is different [11], [12]. The magnetic field sensor can obtain magnetic field strength data on the smartphone, while CSI usually needs to be collected by computer. Second, in the case of multiple APs [8], the amount of data increases exponentially, so does the computational complexity of the corresponding fingerprint location algorithm. Thirdly, what the magnetic field sensor picks up is the three-dimensional magnetic field strength [5], [13]. In the localization process, it is necessary to calculate the pitching angle, heading angle and roll angle of the phone in real time, which will inevitably cause errors in the experimental data and affect the localization accuracy. Finally, the most important thing is how to make the CSI and MFS play their respective advantages in the combination process to achieve organic fusion.

In order to solve the above challenges, we designed CSMS, which organically integrated the advantages of CSI and MFS for localization, and can only rely on several existing Wi-Fi hardware to achieve sub-meter level localization on smartphone. According to [14], we extracted 242 CSI subcarriers from Nexus 5 smartphone at 80M bandwidth, which is much larger than the 30 subcarriers obtained by Intel 5300 wireless network card at 20MHz bandwidth. For purpose of reduce the computation complexity, we analyzed the subcarriers separately and screened out 37 optimal subcarriers. Then we extracted the corresponding characteristic values for each subcarrier and applied dynamic weights to CSI data of different AP during the localization. For the three-dimensional magnetic field strength, for purpose of reduce the error generated by the calculation angle, we only utilized the value of magnetic field strength in CSMS, regardless of the direction [15], [16]. However, the value of MFS is not unique and the error of single point localization is large. In order to solve this problem, we applied continuous magnetic field strength sequence to construct geomagnetic fingerprint, which not only has better uniqueness, but also can more intuitively reflect the distribution of indoor magnetic field. Furthermore, in order to integrate CSI and MFS more organically, we proposed Local Dynamic Time Warping (L-DTW) algorithm to match geomagnetic waveform for tracking and then applied dynamic weights to CSI data of multiple APs according to the tracking results and reduced the localization range. Finally, we matched the fused fingerprint to obtain the localization coordinates according to the proposed Multi-Module Data k-Nearest Neighbor (M-KNN) algorithm.

We conduct extensive experiments in multiple scenarios to validate the performance of CSMS. The results of experimental show that the mean distance error in both scenarios is less than 0.5m which is significantly superior to existing smartphone-based indoor positioning methods.

The main contributions of this work are summarized as follows:

- To the best of our knowledge, it is the first time to propose an indoor localization method combining CSI and MFS based on smartphone. This method overcomes their respective drawbacks and yields great performance that is not achievable by any single sub module alone.
- In view of the long-term stability of geomagnetic signals on the same path indoors, we propose the Local Dynamic Time Warping algorithm for geomagnetic tracking. In addition, according to the tracking location, we propose the Multi-Module Data k-Nearest Neighbor algorithm for fusion fingerprint dynamic weighted comparison.
- We collected extensive data in two scenarios and screened out the optimal subcarriers for experiment. The results of experimental show that the accuracy of CSMS is significantly superior to existing smartphone-based indoor positioning methods.

The remainder of this paper is structured as follows. We present the preliminaries in Section 2, which include the characteristics of geomagnetic localization and the basic principle of CSI, followed by detailed presentation about the architecture of CSMS and related algorithms in Section 3. We implement and evaluate CSMS in Section 4 and conclude this work in Section 5.

II. PRELIMINARIES
A. GEOMAGNETIC LOCALIZATION
Geomagnetic field is the basic physical field of the earth. Any position of the near-earth space has magnetic field strength, and its strength and direction vary with different longitude, latitude and height [15]. Meanwhile, according to magnetic field theory, magnetic materials have an effect on magnetic fields. Nowadays, most buildings are reinforced concrete structures, which bend the geomagnetic field in local space, but are stable in time and have certain uniqueness. Therefore, the magnetic field strength can be applied to indoor positioning [17], [18]. It can be seen from [5] that the strength of indoor geomagnetic field is very stable with time, and the influence of moving objects on the magnetic field is very limited.

In order to verify the influence of device diversity on geomagnetic localization, we exerted three different smartphones to obtain geomagnetic information on the same path. As shown in Fig.1, for different devices, geomagnetic information collected in the same scene has the same trend.

B. CHANNEL STATE INFORMATION (CSI)
With the application of such technologies as orthogonal frequency division multiplexing (OFDM) and multiple-input multiple-output (MIMO) in IEEE 802.11a/n [19] protocol, the channel characteristics between Wi-Fi transceiver devices can be stored as CSI. As a quantitative representation of channel frequency response, CSI can reflect scattering,
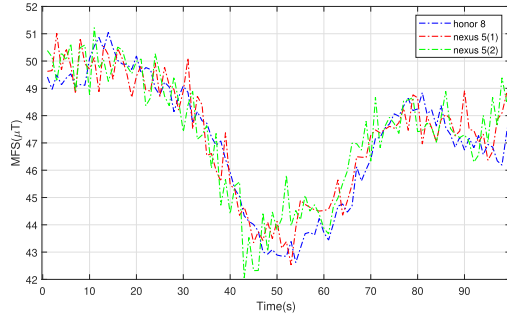
FIGURE 1. Device diversity.

environmental attenuation, power attenuation and other properties in the physical environment.

OFDM [20] is an efficient digital multi-carrier modulation scheme for broadband wireless communication, widely applied in IEEE 802.11a/g/n and WiMAX [21], [22]. It is the core technology of standards such as 3GPP LTE [23]. OFDM divides the channel into several orthogonal sub-channels, and the received signals transmitted through the multipath channel can be represented as:

$$Y = HX + N \tag{1}$$

where Y represents the signal vectors of the receiver and X represents the signal vectors of transmitter. H is the channel information matrix, N is the noise vector, and CSI of each subcarrier can be estimated as:

$$\hat{H} = \frac{Y}{X} \tag{2}$$

where $\hat{H}$ represents the channel frequency response (CFR) of each sub-channel. According to the driver of the underlying hardware device at the receiving end, CSI is divided into multiple subcarrier groups, so the matrix H of CSI can be represented as:

$$H = [H_1, H_2, \cdots H_N] \tag{3}$$

where N is the number of subcarrier groups divided according to the driver. When the channel bandwidth is [20, 40, 80, 160]MHz, N is [56, 114, 242, 484] respectively. H_i is CSI on each subcarrier and is expressed as:

$$H_i = |H_i|\, e^{j \sin(\angle \mathcal{H}_j)} \tag{4}$$

where H_i represents the amplitude and $\angle H_i$ represents the phase of the ith subcarrier.

C. EXTRACTING CSI

Since CSI describes link-layer information, ordinary Wi-Fi hardware will not normally output this information. Therefore, in order to obtain CSI, the following two methods are usually utilized:

(i) Construction of software-defined radio hardware platform (SDR platform) [11]. This method requires a high cost and is not convenient for large-scale promotion.

(ii) By installing the Intel 5300 wireless network card and modifying the drive firmware [12]. Most researchers collect CSI in this way. The limitation of this method is that it can only be applied on computers, and it is difficult to obtain CSI of multiple APs in a short time.

In this study, we utilized the Nexus5 smartphone [14] to collect 242 subcarriers in the 80MHz bandwidth of 44 channel, which is much larger than the 30 subcarriers that Intel 5300 wireless network card can only pick up at 20MHz bandwidth.

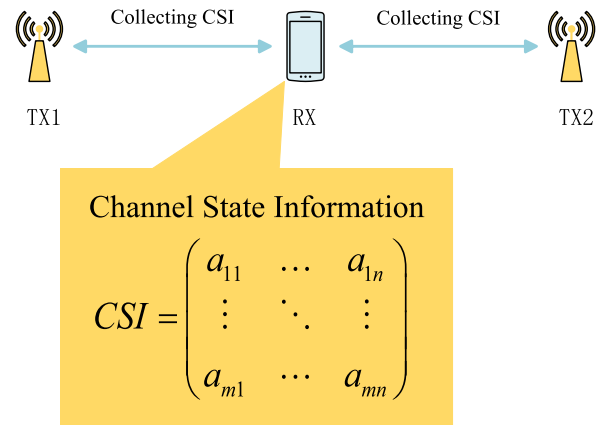As illustrated in Fig.2, n and m represent the number of subcarriers and data packets respectively.

FIGURE 2. Acquisition of channel state information.

III. SYSTEM OVERVIEW

In this section, we first present the system architecture of CSMS, and then introduce the related methodologies in detail.

A. ARCHITECTURE

This study proposes CSMS, an indoor localization method integrating CSI and geomagnetic field strength (GMFS). As illustrated in Fig.3, CSMS consists of a training phase and a localization phase. During the training phase, GMFS and CSI of multiple APs are collected to construct fusion fingerprint map, and geomagnetic data from multiple indoor paths are picked up to construct the geomagnetic fingerprint map. In the localization phase, firstly, the initial positioning coordinates are obtained according to M-KNN algorithm, and then L-DTW algorithm is applied to match the geomagnetic sequence during the movement for tracking. Finally, according to the tracking location, utilize M-KNN algorithm to dynamically weight multi-module data and narrow the positioning range for fingerprint comparison to obtain more accurate localization results.

Various related algorithms proposed in this study will be described in detail below.

B. CONSTRUCTION OF FUSION FINGERPRINT MAP

CSMS utilize Nexus 5 smartphones to collect CSI. Compared with the previous Intel 5300 wireless card, which could
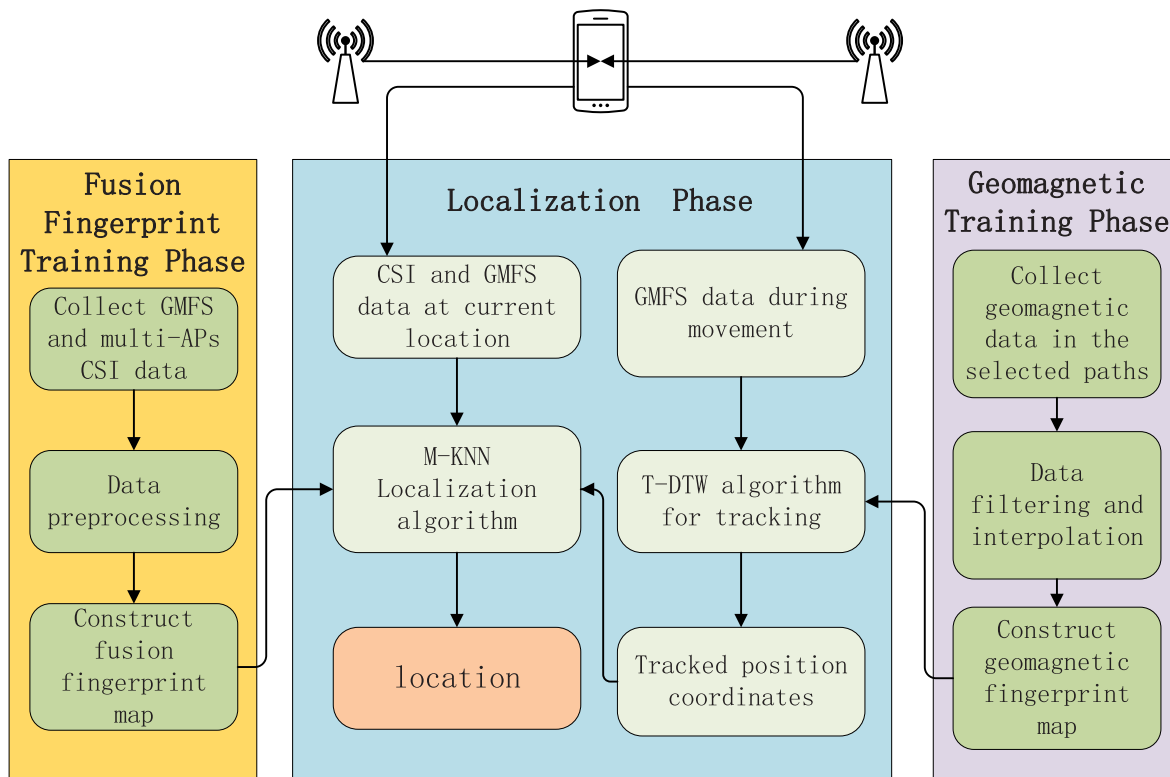
FIGURE 3. Architecture of CSMS.

only obtain 30 subcarriers, the Nexus 5 smartphone can not only acquire 242 subcarriers, but also more conveniently and flexibly.

The magnetic field sensor of the smartphone can collect the three-dimensional magnetic field strength of the current position [24], [25], but the magnetic field data measured at different orientations cannot be matched directly due to the different coordinate systems, so the transformation should be carried out according to the following equations:

$$Mat_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \quad (5)$$

$$Mat_y = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \quad (6)$$

$$Mat_z = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7)$$

$$\begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} = Mat_y \cdot Mat_x \cdot Mat_z \begin{pmatrix} M_x \\ M_y \\ M_z \end{pmatrix} \quad (8)$$

where α, β, γ represent the rotation angles about the X, Y and Z axes respectively, which can be calculated based on the data of the acceleration sensor. $Mat_x, M\ at_y, M\ at_z$ represent the rotation matrix respectively and M_x, M_y, M_z represent the three-dimensional magnetic field strength

obtained by the mobile phone sensor. T_x, T_y, T_z represent the three-dimensional magnetic field strength after conversion.

The steps to build the fusion fingerprint library are as follows:

(1) Select k reference points in indoor scenes, and collect CSI data of n APs at each reference point. Save the data as $FgCSI(k, n)$:

$$FgCSI(k, n) = \begin{pmatrix} fgC_{11} & \cdots & fgC_{1n} \\ \vdots & \ddots & \vdots \\ fgC_{k1} & \cdots & fgC_{kn} \end{pmatrix} \quad (9)$$

where fgC_{ij} represents the CSI data that collects from j-th AP at i-th reference point.

(2) At each reference point, collects three-dimensional magnetic field strength data at m different orientations. Save as $FgMFS(k, m)$:

$$FgMFS(k, m) = \begin{pmatrix} fgM_{11} & \cdots & fgM_{1m} \\ \vdots & \ddots & \vdots \\ fgM_{k1} & \cdots & fgM_{kn} \end{pmatrix} \quad (10)$$

where fgM_{ij} represents the three-dimensional magnetic field strength and corresponding rotation angle that collects from j-th orientation at i-th reference point.

(3) Screen out the optimal subcarriers from each group of CSI data. Then filter and extract the characteristic values as CSI fingerprint. For each three-dimensional

magnetic field strength data $(M_x \ M_y \ M_z)^T$. According to the corresponding rotation angle $(\alpha \ \beta \ \gamma)$, Eq.8 is exerted to calculate the three-dimensional magnetic field strength after the coordinate transformation $(T_x \ T_y \ T_z)^T$. Then, the converted data at different orientations of the same location are averaged as the geomagnetic fingerprint characteristics of the location.

(4) According to the calculation, CSI and geomagnetic fingerprint characteristics of all reference points are saved as a fused fingerprint map $FgF(k, n+1)$:

$$FgF(k, n+1) = \begin{pmatrix} fgfu_{11} & \cdots & fgfu_{1n+1} \\ \vdots & \ddots & \vdots \\ fgfu_{k1} & \cdots & fgfu_{kn+1} \end{pmatrix} \quad (11)$$

where $fgfu_{i1}$ to $fgfu_{in}$ represent the CSI fingerprint characteristic data of n APs at i-th reference point. $fgfu_{in+1}$ represents the geomagnetic fingerprint characteristic data at i-th reference point.

C. MULTI-MODULE DATA K-NEAREST NEIGHBOR(M-KNN)

In this study, we proposed M-KNN algorithm, which is based on the principle of KNN algorithm and can match multi-module data by weighted fusion. The steps of M-KNN are described in detail in algorithm.1.

D. CONSTRUCTION OF GEOMAGNETIC FINGERPRINTS

The magnetic field sensor of the smartphone can collect the three-dimensional magnetic field strength of the current position. If the geomagnetic fingerprint is constructed based on the three-dimensional magnetic field strength, the pitch angle, heading angle, and roll angle of the smartphone need to be calculated in real time. These calculation processes will inevitably cause errors in the experimental data and affect the positioning accuracy. In order to reduce the above errors, this study only employs the strength of the geomagnetic field, regardless of the direction.

The magnetic field strength is calculated from Eq.15, where is the magnetic field strength at any point, M_x, M_y and M_z are the three-dimensional magnetic field strengths in the smartphone coordinate system respectively.

$$M = \sqrt{M_x^2 + M_y^2 + M_z^2} \quad (12)$$

Due to the non-uniqueness of the magnetic field strength value, there is a large error in the single-point positioning. Therefore, this study adopts fast continuous acquisition method to construct indoor geomagnetic fingerprint map. The map is composed of geomagnetic sequences collected on multiple paths in the room continuously, which can reflect the distribution of the indoor magnetic field and has good uniqueness.

Fig.4 shows the acquisition of geomagnetic information on a path indoors continuously.

Algorithm 1 Multi-Module Data k-Nearest Neighbor

Input: CSI collected from n APs: $mat_1, mat_2, \cdots mat_n$, Three-dimensional magnetic field strength after conversion: mat_{n+1}, The coordinates tracked by the geomagnetic field: (x_0, y_0);

Output: Positioning coordinate (x_1, y_1);

1: **for** i = 1 : n **do**
2:     Calculate the distance to i-th AP $dist_i$ based on the position coordinates (x_0, y_0);
3:     **if** $(x_0, y_0)! = (-1, -1)$ **then**
4:         According to the dynamic weight formula $P = -0.015 \times dist + 0.95$ proposed by us, the weight P_i corresponding to the i-th AP is calculated;
5:     **else**
6:         The corresponding weight of i-th AP is 1;
7:     **end if**
8: **end for**
9: Take the Hadamard product of $(P_1, P_2, \cdots P_n, 1)$ and $(mat_1, mat_2, \cdots mat_{n+1})$;
10: $finger_1 (1, n+1) = (P_1, P_2, \cdots P_n, 1) * (mat_1, mat_2, \cdots mat_{n+1})$;
11: **if** $(x_0, y_0)! = (-1, -1)$ **then**
12:     Take the reference point within 5 meters from (x_0, y_0) as the sub-fingerprint map $FingerF_1 (k, n+1)$;
13: **else**
14:     Take the original fingerprint map as $FingerF_1 (k, n+1)$;
15: **end if**
16: count = size$(FingerF_1)$;
17: **for** j = 1 : count **do**
18:     Calculate the Euclidean distance D_j from $finger_1 (1, n+1)$ to the j-th element of $FingerF_1 (k, n+1)$;
19: **end for**
20: Sort $(D_1, D_2, \cdots D_{count})$ from small to large;
21: $Sequence_D = Sort(D_1, D_2, \cdots D_{count})$;
22: Take the reference point coordinate corresponding to the first element in $Sequence_D$ as the positioning coordinate (x_1, y_1);
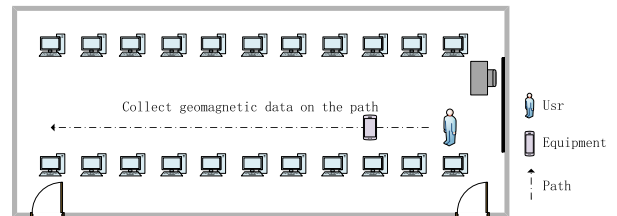23: **return** (x_1, y_1);

FIGURE 4. Acquisition of geomagnetic data.

E. LOCAL DYNAMIC TIME WARPING (L-DTW)

When the traditional DTW algorithm [26], [27] matches two time series, it can only calculate the similarity between

the two sequences from the start to the end, and cannot perform local matching. Therefore, in order to match the MFS sequence locally, this study proposes L-DTW algorithm according to the characteristics of geomagnetic information. By improving the DTW algorithm, L-DTW achieves local matching of time series with different lengths.

Algorithm.2 introduces the L-DTW algorithm in detail.

Algorithm 2 Local Dynamic Time Warping

Input: Test fingerprint sequence $X(n) = x_1, x_2, \cdots x_n$ and reference fingerprint sequence $Y(m) = y_1, y_2, \cdots y_m$;

Output: The similarity SIM_{XY} of the final local match (the smaller the value, the more similar) and the final position END_Y matched on sequence $Y(m)$;

1: n = size(X);
2: m = size(Y);
3: $sqrty = \lfloor \sqrt{m} \rfloor$;
4: Starting from the first point y_1 of Y, find K initial positions $p_1, p_2, \cdots p_k$ on Y spaced by $sqrty$:
5: $P(k) = p_1, p_2, \cdots p_k$;
6: Construct $n*m$ matrix $d(n, m)$, where $d(i, j) = |x_i - y_j|$;
7: $SIM_{XY} = realmax$;
8: **for** i = i : k **do**
9:     Take the i-th position p_i from $P(k)$.
10:     Construct $n * m$ matrix $D(n, m)$, and assign the first row and the p_i column of $d(n, m)$ to $D(n, m)$:
11:     $D(1, :) = d(1, :)$;
12:     $D(:, p_i) = d(:, p_i)$;
13:     **for** j = p_i + 1 : p_i + 2n **do**
14:         **for** s = 2 : n **do**
15:             $D(s, j) = \min(D(s - 1, j), D(s - 1, j - 1), D(s, j - 1)) + d(s, j)$
16:         **end for**
17:     **end for**
18:     Compare the minimum $D(n, min) = minD$ in $D(n, p_i : p_i + 2n)$ with SIM_{XY};
19:     **if** $minD < SIM_{XY}$ **then**
20:         $SIM_{XY} = minD$;
21:         $END_Y = min$;
22:     **end if**
23: **end for**
24: **return** (SIM_{XY}, END_Y);

IV. PERFORMANCE EVALUATION

In this section, we introduce our experimental scenarios and verify the feasibility of geomagnetic tracking firstly. Afterwards, we will show the performance of CSMS by comparing with other two CSI-based fingerprint localization methods.

A. EXPERIMENTAL SCENARIOS

In the experiment, two TL-WDR5610 routers manufactured by TP-LINK as the Access Points (APs). A Nexus 5 smartphone served as the receiver object, which equipped with a BCM4339 Wi-Fi chip. The two different scenarios for the experiment are as follows:
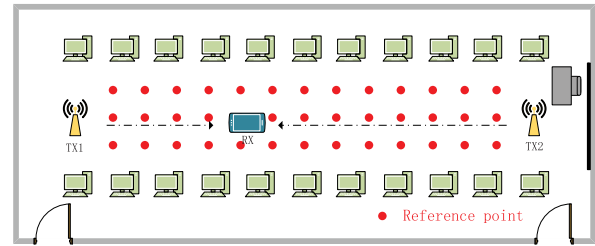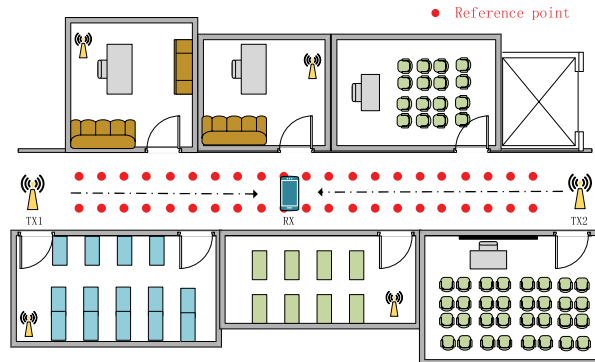
FIGURE 5. Research laboratory.

FIGURE 6. Rectangular corridor.

(i) Research Laboratory
First, we experiment in an $8m \times 20m$ research laboratory covering by two APs as shown in Fig.5. Two APs were fixed at the opposite ends of the experimental scenario and 3×13 reference points were selected to collect CSI, among which every two adjacent reference points were 1.2 meters apart.

(ii) Rectangular Corridor
Second, we conducted experiments in a rectangular corridor with multiple offices, which is $2.4m \times 30m$ covering corridors, rooms and classrooms as shown in Fig.6. In this scenario, there were also two APs were fixed at the opposite ends and 2×21 reference points were selected to collect CSI.

In each experimental scenario, we chose multiple straight paths for picking up geomagnetic data.

B. GEOMAGNETIC TRACKING

In order to verify the feasibility of L-DTW algorithm for geomagnetic tracking, we divide the experiment into training phase and tracking phase. Geomagnetic fingerprints are created in the training phase, and during the tracking phase, we build a test data set to test the performance of L-DTW.

1) TRAINING PHASE

In this phase, we utilized the method of fast continuous acquisition to obtain MFS sequences. For each path selected, we collected three round trips at a constant speed. Then in the experiment, we found the positioning accuracy was not

high due to the low sampling frequency of the geomagnetic sensor of Nexus 5 smartphone. To solve this problem, we first performed outlier processing on the data, then Fourier interpolation was be exerted to expand the data and the expanded data was employed to construct a geomagnetic fingerprint map.

2) TRACKING PHASE

In the tracking phase, we collected four sets of geomagnetic sequences with the path length of 3m, 4m, 5m and 6m on the indoor linear path respectively, and recorded the corresponding start and end positions of each sequence. Then performed outlier processing and data expansion. Finally, we built a test data set based on the processed data for testing the performance of geomagnetic tracking.

We utilized the L-DTW algorithm to match the geomagnetic sequence in the test set with the fingerprint sequence in the geomagnetic fingerprint map successively. Then, according to the matching results, we applied the KNN algorithm [28] to find the path with the best matching similarity for each test sequence. Moreover, according to the last position matched in the path fingerprint sequence, the indoor tracking position of the test sequence is determined.

As we can see from Tab.1, the experimental results show that the longer the path length of the geomagnetic sequence in the test set, the higher the positioning accuracy. When the acquisition path length is 6 meters, the average positioning error of this method is 1.1887 meters.

TABLE 1. Error analysis of geomagnetic tracking.

Path length(m)	Maximum error(m)	Minimum error(m)	Average error(m)	Standard deviation
3	7.4192	0.3185	1.5899	1.5789
4	6.2168	0.2979	1.4374	1.3506
5	4.9754	0.2880	1.3157	1.1029
6	4.7768	0.2383	1.1887	0.8355

Fig.7 shows the CDF of the positioning errors at the acquisition paths of 3m, 4m, 5m and 6m. We can see that the accuracy of the positioning error within 1.5 meters for the
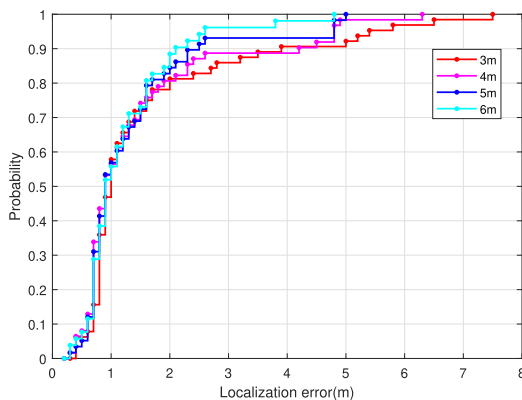

FIGURE 7. Tracking accuracy of L-DTW.

four types of data is relatively close, all of which are about 0.72. As the acquisition path length increases, the maximum positioning error decreases gradually. With a collection path length of 6 meters, the positioning accuracy within 2.5 meters of the positioning error can reach 0.94.

C. CSMS LOCALIZATION

The experiment of CSMS consists of Calibration Phase and Positioning Phase. In the Calibration Phase, we process the multi-module data and build a fusion fingerprint map. In the Positioning Phase, we utilize the M-KNN algorithm for fusion fingerprint dynamic weighted comparison.

1) CALIBRATION PHASE

In this phase, we collected CSI and MFS data through smartphone to build a fusion fingerprint map. Through analyzing CSI, we found that different subcarriers were affected differently by environmental changes. In order to make the localization results more accurate, we analyzed 242 subcarriers separately, from which we screened out 37 optimal subcarriers that were significantly affected by the environment, and employed these subcarriers for experiments.

Based on the reference points selected in two experimental scenarios, as shown in Fig.8, we first collected 3000 CSI data packets from two APs at any reference point at 44-channel and 80M bandwidth, each packet contains 242 subcarriers. Then we screened out the 37 optimal subcarriers as illustrated in Fig.9 and get two 3000 * 37 CSI matrices. Furthermore, hampel filter was exerted to remove outliers for each column of the matrix, and then kalman filter is performed, as shown in Fig.10. Finally, Fig.11 shows that we calculated the average value and the values of 1/4, 1/2, and 3/4 quantiles for each column of the filtered matrix to obtain two 4 * 37 characteristic values matrices $fingerfu_{i1}$ and $fingerfu_{i2}$.
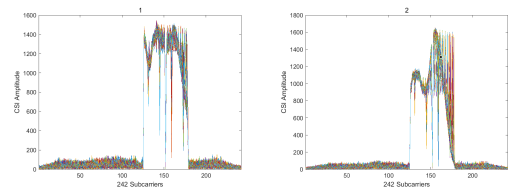

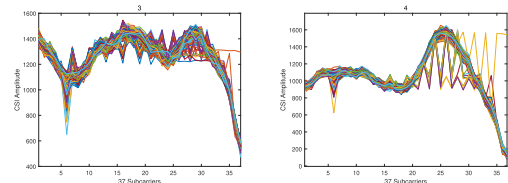**FIGURE 8.** The 242 subcarriers collected at two reference points.

FIGURE 9. The 37 optimal subcarriers of two reference points.

For any reference point, we collected three-dimensional magnetic field strength data from six different orientations, and then removed the outliers and averaged each of the
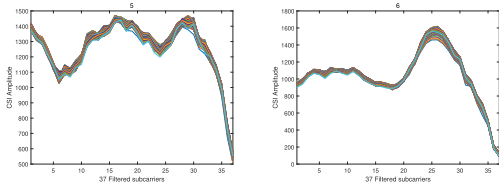
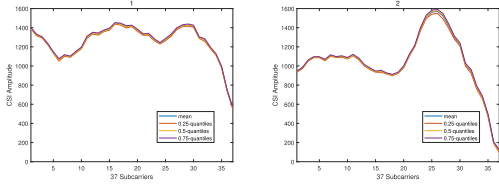**FIGURE 10. The 37 optimal subcarriers after filtering.**

FIGURE 11. Characteristic values extracted from 37 optimal subcarriers.

three-axis data separately. Next, we combined the rotation angle corresponding to each orientation and calculated the three-dimensional magnetic field strength after coordinate conversion by Eq.8. Then we averaged the data converted from six different orientations to obtain the geomagnetic fingerprint feature matrix $fingerfu_{i3}$.

Finally, we utilized the feature matrix of all reference points to build a multi-module fusion fingerprint database $FgF(k, 3)$:

$$FgF(k, 3) = \begin{pmatrix} fgfu_{11} & \dots & fgfu_{13} \\ \vdots & \ddots & \vdots \\ fgfu_{k1} & \dots & fgfu_{k3} \end{pmatrix} \quad (13)$$

where k represents the number of reference points.

2) POSITIONING PHASE

For location estimation phase, we collected multiple sets of MFS and CSI data at the start and end positions of the 6m test geomagnetic sequence in Section IV-B to form the experimental test data set. For the CSI of each start and end point in the data set, we performed corresponding filtering and screened out the 37 optimal subcarriers to obtain two 37-column start-point test matrices $matA_0$ and $matB_0$ and two 37-column end-point test matrices $matA_1$ and $matB_1$. For the MFS in the data set, we calculated the transformed three-axis magnetic field strength according to the corresponding rotation angle, and obtained the starting and ending test matrices $matC_0$ and $matC_1$.

According to the M-KNN algorithm, we matched the starting point test matrix ($matA_0$, $matB_0$, $matC_0$) with each row of the fingerprint database $FingerF(k, 3)$ to find the corresponding positioning coordinate matrix CD_0. Then we correspondingly narrow the range that needs to be matched in the geomagnetic fingerprint map, and use the L-DTW algorithm to find the end position CD_1 of each 6m sequence. Finally, we take the coordinates in CD_1 as input and use the M-KNN algorithm to perform weighted matching on the data to find the final localization result.

3) PERFORMANCE EVALUATION

In our experiments, we compared the performance of CSMS with the CSI-based fingerprint localization of single AP and the CSI-based fingerprint localization of double APs. Fig.12 shows the cumulative distribution function (CDF) of positioning errors in the laboratory. There are 39 points in the laboratory for fingerprint collection. As we can see, the positioning error of CSMS methods is much smaller than the other two methods and the accuracy of CSMS within 2m can reach 99%.
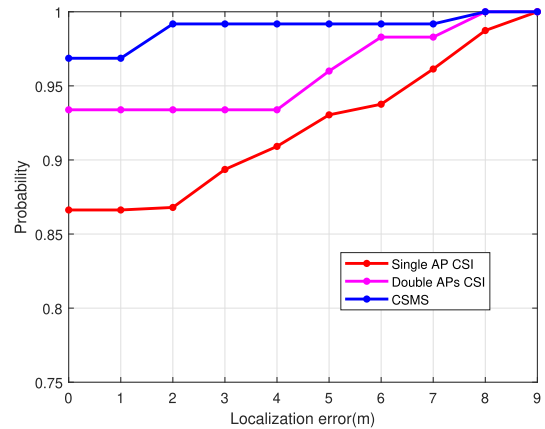

FIGURE 12. CDF of localization error in the laboratory.

Unlike the first scenario that 2 APs and smartphone are placed in the airtight room, we also conducted experiments in the hallway. Fig.13 illustrates the cumulative distribution of localization errors across 42 positions in the corridor. We can easily observe that the accuracy within 3m by single AP is much smaller than the other two methods and the positioning accuracy of the three methods above 4m is similar. However, the positioning accuracy of CSMS method is still slightly higher.
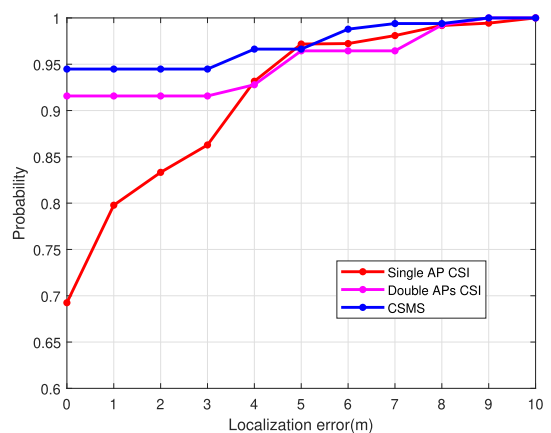

FIGURE 13. CDF of localization error in the corridor.

We evaluate the accuracy of CSMS method and compare it with two other fingerprint-positioning methods, Well-known CSI-based positioning. Fig.14 presents the accuracy of three
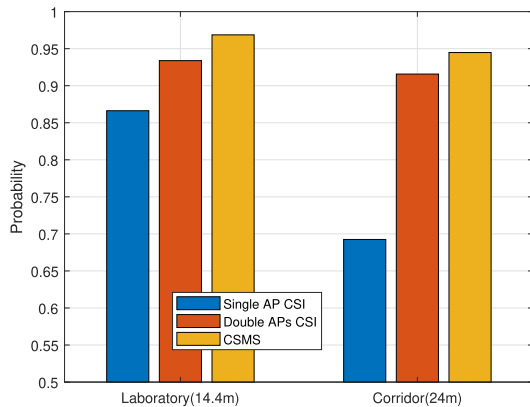
FIGURE 14. Accuracy of three positioning methods.

different methods in two environments. As shown in the figure, the positioning accuracy of CSMS is higher than the other two methods in both scenarios and the accuracy of CSMS in the laboratory can reach 96.86%, which is 3.73% higher than the double APs case. In addition, we compare the mean distance error of three methods in two scenarios. As illustrated in Fig.15, the mean distance error of CSMS in the laboratory is significantly smaller than the other two methods, which can reach 0.097m and it can reach 0.381m even in the corridor. We can find that the CSMS method reflects a preferable property since the organic integration of CSI and MFS is beneficial to improve the accuracy of location fingerprinting.
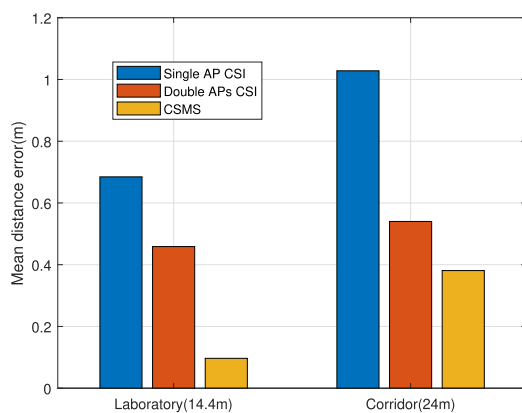

FIGURE 15. Mean distance error of three positioning methods.

V. CONCLUSION

In this paper, we present CSMS, a smartphone-based sub-meter accuracy indoor localization method that integrates CSI and MFS. By fusing data from multiple sub modules, CSMS successfully overcomes their respective drawbacks and yields great performance that is not achievable by any single sub module alone. We conduct extensive experiments in multiple scenarios to validate the performance of CSMS. The results of experimental show that the mean distance error in both scenarios is less than 0.5m which is significantly

superior to existing smartphone-based indoor positioning methods.

REFERENCES

[1] Y. Alvarez and F. Las Heras, "ZigBee-based sensor network for indoor location and tracking applications," *IEEE Latin Amer. Trans.*, vol. 14, no. 7, pp. 3208–3214, Jul. 2016.

[2] J. D. Boston and B. J. Smith, "Multi-floor and multi-building ultra-wideband (UWB) location device, system, and method," U.S. Patent 10 064 012, Aug. 28, 2018.

[3] K. Shoarinejad, M. Soltan, and M. Moshfeghi, "Rfid location systems and methods," U.S. Patent 8 294 554, Oct. 23, 2012.

[4] H.-S. Kim and J.-S. Choi, "Advanced indoor localization using ultrasonic sensor and digital compass," in *Proc. Int. Conf. Control, Autom. Syst.*, Oct. 2008, pp. 223–226.

[5] Y. Shu, K. G. Shin, T. He, and J. Chen, "Last-mile navigation using smartphones," in *Proc. 21st Annu. Int. Conf. Mobile Computing Netw. (MobiCom)*, 2015, pp. 512–524.

[6] T. Li, H. Wang, Y. Shao, and Q. Niu, "Channel state information–based multi-level fingerprinting for indoor localization with deep learning," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 10, Oct. 2018, Art. no. 155014771880671.

[7] X. Yang, P. Chen, S. Gao, and Q. Niu, "CSI-based low-duty-cycle wireless multimedia sensor network for security monitoring," *Electron. Lett.*, vol. 54, no. 5, pp. 323–324, Mar. 2018.

[8] J. Xiao, K. Wu, Y. Yi, and L. M. Ni, "FIFS: Fine-grained indoor fingerprinting system," in *Proc. 21st Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 2012, pp. 1–7.

[9] C. Wu, Z. Yang, C. Xiao, C. Yang, Y. Liu, and M. Liu, "Static power of mobile devices: Self-updating radio maps for wireless indoor localization," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 2497–2505.

[10] C. Wu, J. Xu, Z. Yang, N. D. Lane, and Z. Yin, "Gain without pain: Accurate WiFi-based Localization using fingerprint spatial gradient," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, no. 2, pp. 1–19, Jun. 2017.

[11] Q. Wang, D. Wu, J. Eilert, and D. Liu, "Cost analysis of channel estimation in MIMO-OFDM for software defined radio," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2008, pp. 935–939.

[12] X. Wang, X. Wang, and S. Mao, "CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[13] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, and M. Wiseman, "Indoor location sensing using geo-magnetism," in *Proc. 9th Int. Conf. Mobile Syst., Appl., Services (MobiSys)*, 2011, pp. 141–154.

[14] M. Schulz, J. Link, F. Gringoli, and M. Hollick, "Shadow Wi-Fi: Teaching smartphones to transmit raw signals and to extract channel state information to implement practical covert channels over Wi-Fi," in *Proc. 16th Annu. Int. Conf. Mobile Syst., Appl., Services (MobiSys)*, 2018, pp. 256–268.

[15] Z. Song, J. Zhang, W. Zhu, and X. Xi, "The vector matching method in geomagnetic aiding navigation," *Sensors*, vol. 16, no. 7, p. 1120, Jul. 2016.

[16] W.-Q. Zhu, Z.-G. Song, X.-L. Xi, and J.-S. Zhang, "The improved ICCP algorithm based on Procrustes analysis for geomagnetic matching navigation," in *Proc. Design, Manuf. Mechatronics*, Nov. 2015, pp. 623–631.

[17] G. J. Shen, C. Zhao, and F. Zhao, "Indoor location-finding using magnetic field anomalies," U.S. Patent 9 326 103, Apr. 26, 2016.

[18] W. Storms, J. Shockley, and J. Raquet, "Magnetic field navigation in an indoor environment," in *Proc. Ubiquitous Positioning Indoor Navigat. Location Based Service*, Oct. 2010, pp. 1–10.

[19] R. Banerjea and H. Ramamurthy, "Coexistence mechanism for WiMAX and IEEE 802.11," U.S. Patent 8 134 988, Mar. 13, 2012.

[20] R. V. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Norwood, MA, USA: Artech House, 2000.

[21] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals WiMAX: Understanding Broadband Wireless Networking*. London, U.K.: Pearson, 2007.

[22] S. A. Ahson and M. Ilyas, *WiMAX: Technologies, Performance Analysis, and QoS*. Boca Raton, FL, USA: CRC Press, 2018.

[23] K. Fazel and S. Kaiser, *Multi-Carrier Spread Spectrum Systems: From OFDM MC-CDMA to LTE WiMAX*. Hoboken, NJ, USA: Wiley, 2008.

[24] F. Memarzanjany and M. Zheng, "Tracking position and movement using a magnetic field," U.S. Patent 10 151 606, Dec. 11, 2018.

[25] A. Govari, A. C. Altmann, and Y. Ephrath, "Position tracking using quasi-dc magnetic fields," U.S. Patent 9 295 529, Mar. 29, 2016.

[26] A. Mueen and E. Keogh, "Extracting optimal performance from dynamic time warping," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2016, pp. 2129–2130.

[27] R. J. Kate, "Using dynamic time warping distances as features for improved time series classification," *Data Min. Knowl. Disc.*, vol. 30, no. 2, pp. 283–312, Mar. 2016.

[28] Z. Deng, X. Zhu, D. Cheng, M. Zong, and S. Zhang, "Efficient kNN classification algorithm for big data," *Neurocomputing*, vol. 195, pp. 143–148, Jun. 2016.

PEIHAO LI was born in China, in 1994. He is currently pursuing the M.D. degree with the Department of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China. He is also researching the indoor localization and sensor networks with the China University of Mining and Technology.

XU YANG was born in China, in 1992. He is currently pursuing the Ph.D. degree with the Department of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China. He is also researching the distributed measurement systems, big data, and sensor networks with the China University of Mining and Technology.

YUQING YIN was born in China, in 1992. She is currently pursuing the Ph.D. degree with the Department of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China. She is also researching the distributed measurement systems and sensor networks with the China University of Mining and Technology.

SHOUWAN GAO was born in China, in 1982. She received the Ph.D. degree from the Department of Computer Science and Technology, Ocean University of China, Qingdao, China. She is currently an Associate Professor with the Department of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China. Her main research interests are sensor networks, distributed measurement systems, Ocean observation networks, and data modeling.

QIANG NIU was born in China, in 1974. He received the Ph.D. degree from the Department of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China. He is currently a Professor and the Associate Dean of the Department of Computer Science and Technology, China University of Mining and Technology. He is also the Deputy Director of the Mine Information Engineering Research Center, Chinese Coal Industry. His main research interests are intelligent information processing, artificial intelligence and pattern recognition, machine learning and data mining, and other aspects of research. He has published more than ten academic articles retrieved by SCI.

• • •